

Een robot is geen vriend maar een slaaf — dat moet zo blijven'

We staan aan de vooravond van de opkomst van de zelfdenkende machine. Belangrijkste taak voor de mens: de controle niet uit handen geven.



Een robot in de houding van 'De denker' van de beeldhouwer Rodin. Wie robots menselijke eigenschappen meegeeft, riskeert dat we ze als vriend gaan zien.

Een techneuten reageren soms wat vermoeid als je ze het 'trolleyprobleem' voorlegt. Het is een bekend voorbeeld van een moreel dilemma dat je met goed fatsoen niet aan een computer kunt overlaten: een tram ('trolley') denderd in volle vaart op een groep mensen af. Een ramp is alleen te voorkomen door snel een wissel om te gooien, waardoor maar één persoon wordt overreden. Wie red je? En wie offer je op?

Voor mensen is dit al een onmogelijk dilemma, laat staan voor kunstmatige intelligentie. Een zelfrijdende auto heeft geen moreel kompas als hij moet kiezen tussen het leven van een voorbijganger óf dat van zijn passagier. 'Twee jaar terug had hier nog bijna niemand het over, maar slimme machines worden steeds sneller realiteit', zegt Frank van Harmelen, hoogleraar knowledge representation & reasoning aan de Vrije Universiteit. 'Kunstmatige intelligentie kan zulke grote gevolgen hebben voor onze maatschappij dat we het hier nu al over moeten hebben.'

Van Harmelen is de drijvende kracht achter de 22e editie van de European Convention on Artificial Intelligence (ECAI), die begin september plaatsvond in Den Haag. Daar gingen wetenschappers uit de hele wereld met elkaar in discussie over ethische vraagstukken en kunstmatige intelligentie (KI). Discussie was er echter weinig: allen waren het met elkaar eens dat KI een interessante ontwikkeling is, maar de mens moet wel opletten om niet de controle uit handen te geven. De mens moet 'in the loop' blijven.

'Enige uitzonderingen daargelaten heeft de mens een diep moreel besef. Maar een zelfdenkende robot heeft dat besef niet', waarschuwt Michael Bratman, hoogleraar en filosoof aan de Amerikaanse Stanford Universiteit. 'We verwachten van soldaten dat zij weloverwogen keuzes maken op het slagveld. Maar een machine ziet alleen de taak die hij moet uitvoeren. Je kunt wel voorwaarden programmeren waar hij rekening mee moet houden, maar we zijn nog lang niet zo ver dat we op dit vlak de nuances van het menselijk brein kunnen nabootsen - als ons dat ooit al gaat lukken.'

Om het wat concreter te maken: je kunt wel programmeren dat een drone een terrorist niet opblaast als die kinderen gijzelt. Maar wat doe je als de terrorist tijdens die gijzeling op een knop dreigt te drukken die een atoombom laat afgaan? Volgens Van Harmelen is dit het moment om het hierover te hebben. 'We gaan ons vaak pas druk maken over de gevolgen van een nieuwe technologie als hij voor de deur staat. Dat kunnen we ons met kunstmatige intelligentie niet veroorloven.'



Bij de algoritmen die aandelen kopen, snapt de programmeur soms niet meer hoe het precies werkt

Nog even en KI duikt overal op. In de vorm van zelfrijdende auto's, maar ook zorgrobots, avatars in het onderwijs en geautomatiseerde beurshandelaren staan binnen een paar jaar op de stoep. Of zijn er nu al. Steeds meer van die systemen zijn zelflerend: ze zien waar het beter kan en passen hun handelingen aan.

'Er zijn nu al algoritmes geschreven die het beste moment kiezen om aandelen te kopen of verkopen. Maar in sommige gevallen begrijpt zelfs de programmeur niet meer precies hoe het algoritme tot zijn beslissing komt', weet Ibo van de Poel, hoogleraar ethiek en technologie aan de

TU Delft. Zulke systemen hebben maar één doel: zo veel mogelijk verdienen aan aandelen. Het grote plaatje zien ze niet.'

Robot die zelf beslissingen nemen zonder dat de makers ervan weten hoe ze tot die beslissingen komen is iets wat we niet moeten willen. Dat schreven Tesla-oprichter Elon Musk, natuurkundige Stephen Hawking en Apple-oprichter Steve Wozniak samen met duizend hoog aangeschreven robotonderzoekers eerder dit jaar in een brandbrief. Daarin pleitten ze voor een verbod op autonome wapens, zoals zelfdenkende drones die op eigen houtje doelwitten kunnen kiezen en uitschakelen.



Onderzoekers willen een verbod op autonome wapens, zoals drones die zelf doelen kiezen en uitschakelen

'Zulke complexe beslissingen zouden altijd bij een mens moeten liggen', vindt Miriam Struik van de vredesbeweging PAX, dat net als Musk en consorten pleit voor een verbod op autonome wapens. 'We zijn niet tegen het gebruik van álle drones, maar er zijn grenzen aan wat we robots zelf moeten laten beslissen.'

Menselijke waarden en normen overbrengen naar slimme machines wordt een helse taak, vindt ook Joanna Bryson, onderzoeker aan de Bath University in Groot-Brittannië. En we moeten dat ook niet willen. 'Als we robots menselijke eigenschappen meegeven, lopen we het risico dat we ze als vriend gaan zien. Maar een robot is geen vriend. Hij kan zich ook niet inleven in wat het is om mens te zijn. Hoe goed je hem ook programmeert. Robots zijn slaven en moeten dat blijven.'

Nick Kivits is freelancewetenschapsjournalist.

Andere bronnen

Computerbeurs

In de VS en Europa wordt zo'n 75 van de waarde van alle aandelen gekocht en verkocht door computerprogramma's. Vooral grote spelers gebruiken dergelijke algoritmes, maar ook kleinere en onafhankelijke handelaren grijpen steeds vaker naar software. Dat is mede te danken aan gecrowdsourcede projecten als Quantopian en QuantConnect die zelf algoritmes laten ontwikkelen.



.....

Betere gevechtspiloten

Sommige kunstmatige intelligentie is nu al beter in luchtgevechten dan mensen. Het door de Universiteit van Cincinnati ontwikkelde ALPHA was in tests de gepensioneerde Amerikaanse luchtmachtgeneraal Gene Lee te slim af. Lee, die al sinds de jaren 1980 KI-systemen test, was onder de indruk: 'Het leek precies te weten wat ik van plan was.'



Bug van -440 miljoen

Computeralgoritmes aandelen laten verhandelen is niet zonder risico's. In 2012 zag de investeringsgroep Knight Capital 440 miljoen dollar (ongeveer vier keer de netto omzet van 2011) van zijn kapitaal verdampen door een fout van een computersysteem. De bug zorgde ervoor dat het systeem aandelen hoog inkocht en laag verkocht.

